



Good Performance for Bad Days

or: are we measuring the right thing?

Marc Brooker

VP/Distinguished Engineer
Amazon Web Services

Who Am I?

- 16 years at AWS
 - 16 years of carrying a pager
- Worked on some of the largest cloud systems
 - EC2, EBS, Lambda, Aurora, Aurora DSQL, Bedrock, etc.
- Interested in system performance and stability
- Read about 4000 post-mortems
- <https://brooker.co.za/blog/>

Hypothesis 1:

In large-scale systems, system performance is the single largest contributor to system availability.

Hypothesis 1b:

In large-scale systems, **unpredictable** system performance is the single largest contributor to system **un**availability.

Hypothesis 2:

Performance evaluation's primary value is helping system designers make good decisions.

Hypothesis 3:

We're not doing a great job of that.

GOOD PERFORMANCE FOR BAD DAYS

Why?

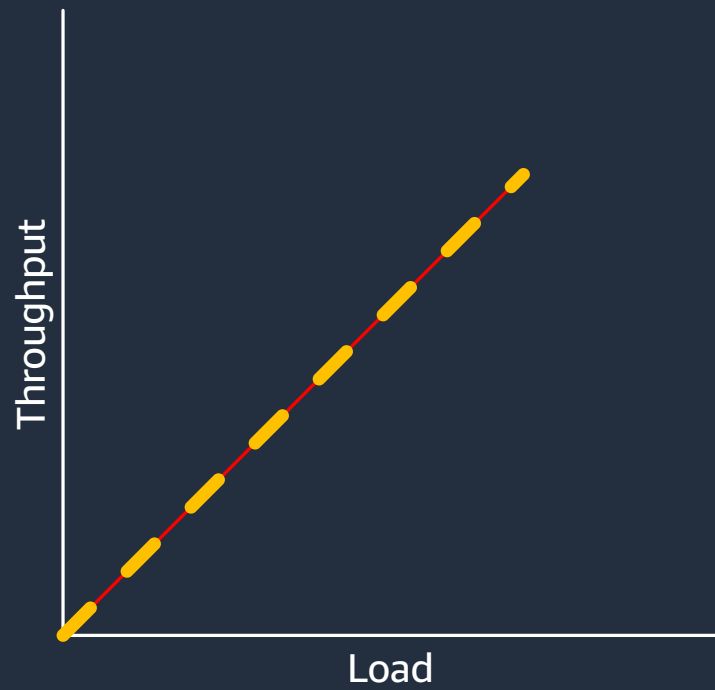


© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

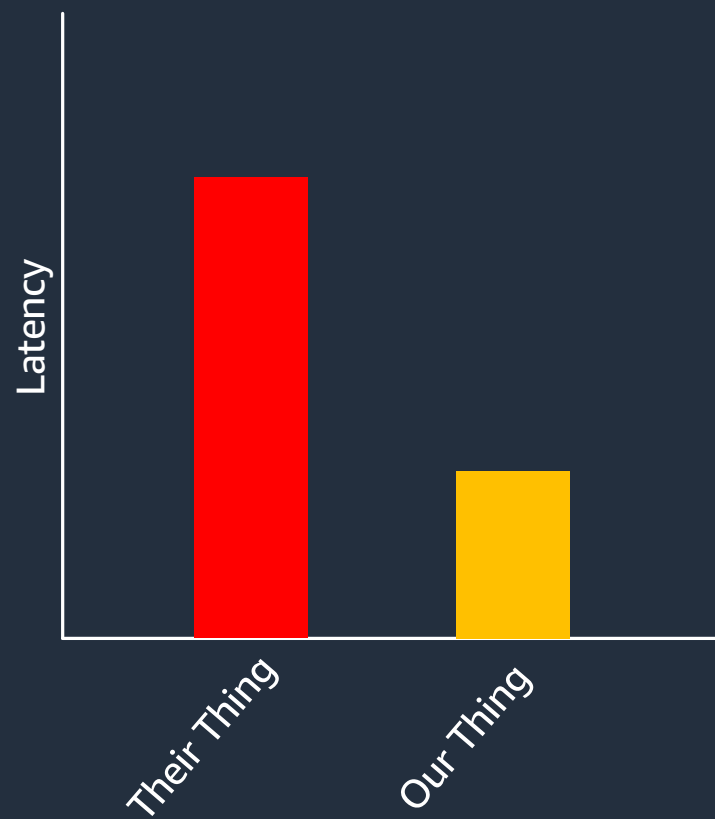
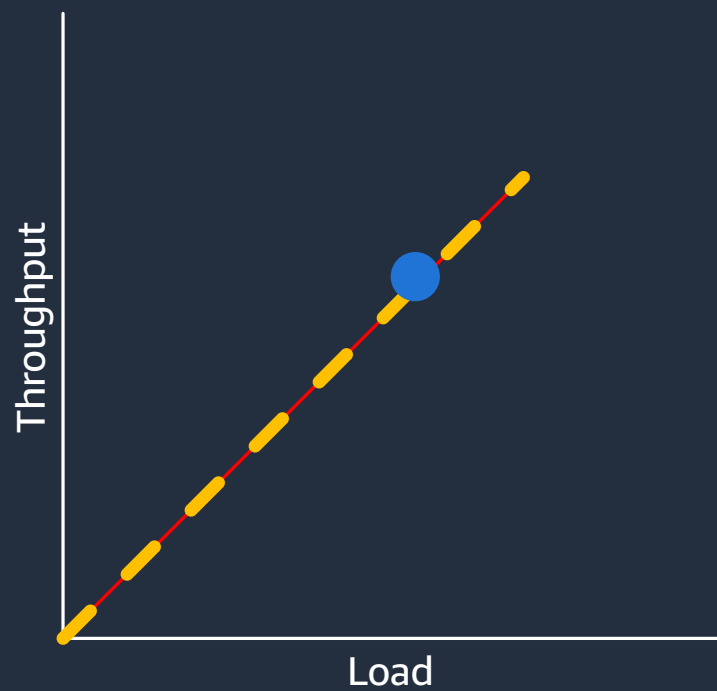
Why 1: Insufficient Attention to Overload Behavior

Not paying enough attention to the behavior of systems under excess or unexpected load.

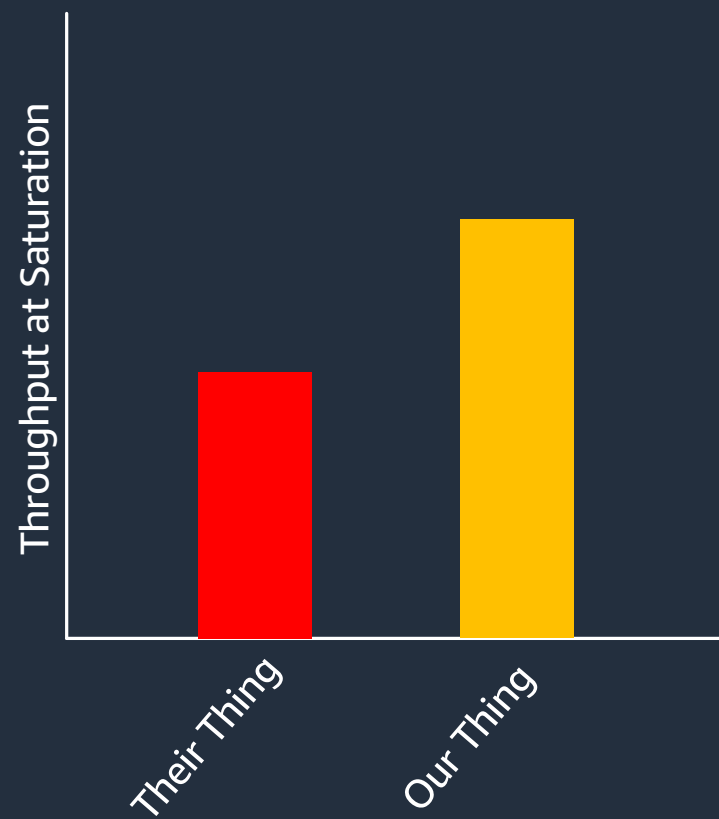
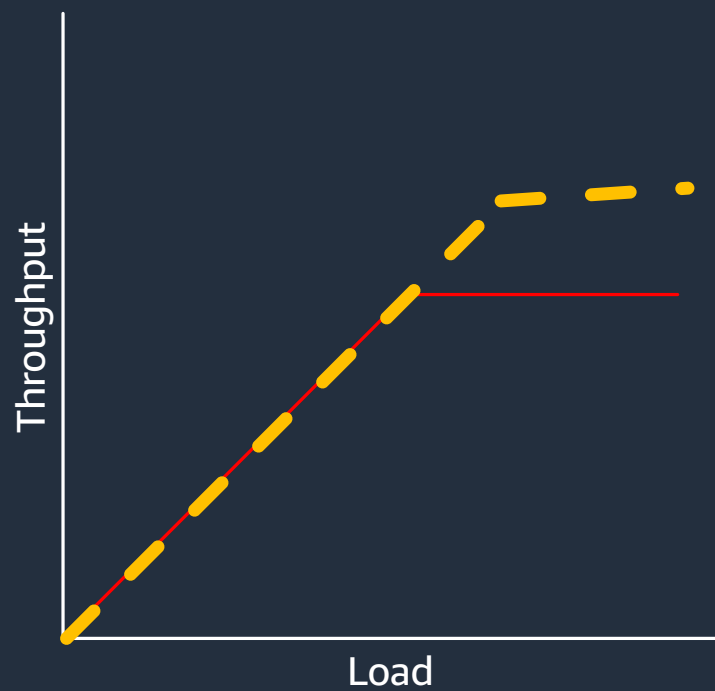
GOOD PERFORMANCE FOR BAD DAYS



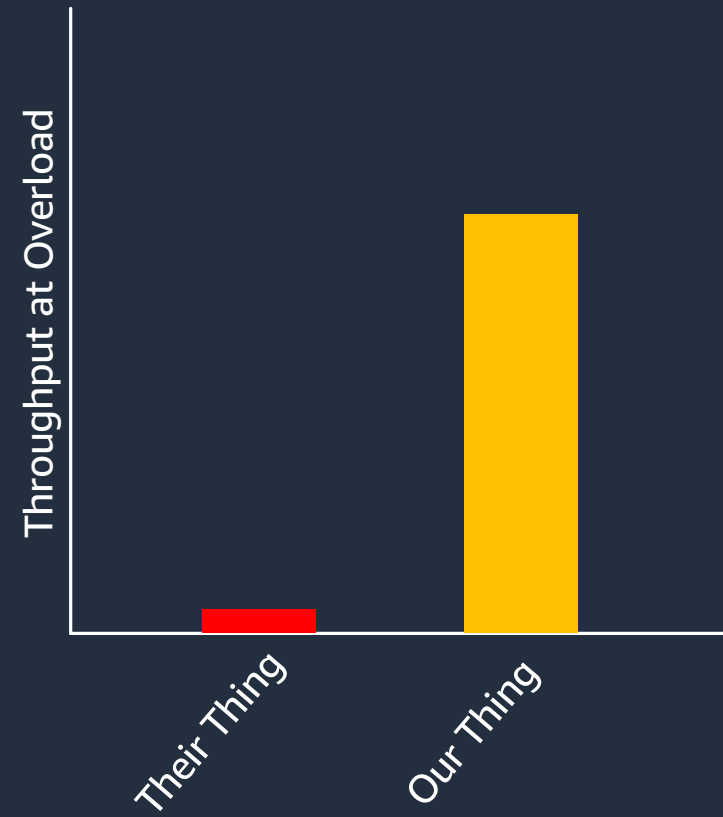
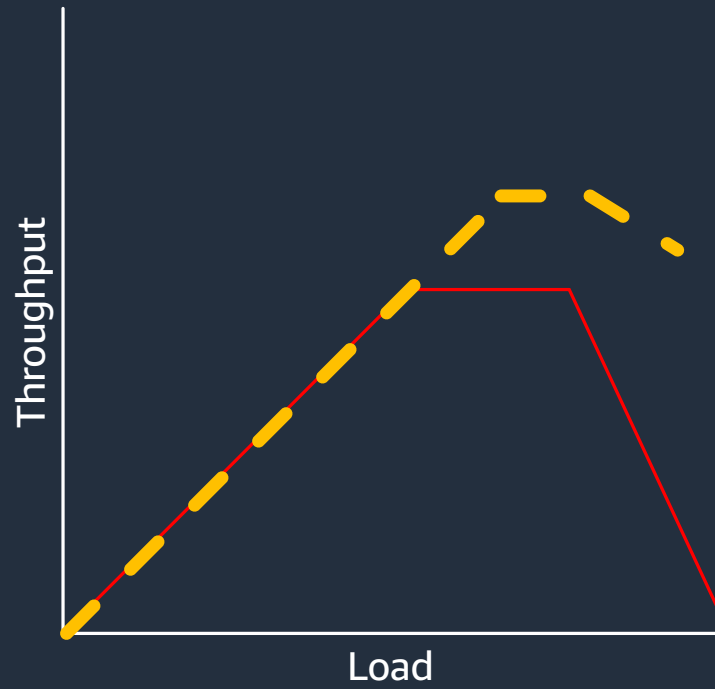
GOOD PERFORMANCE FOR BAD DAYS



GOOD PERFORMANCE FOR BAD DAYS



GOOD PERFORMANCE FOR BAD DAYS



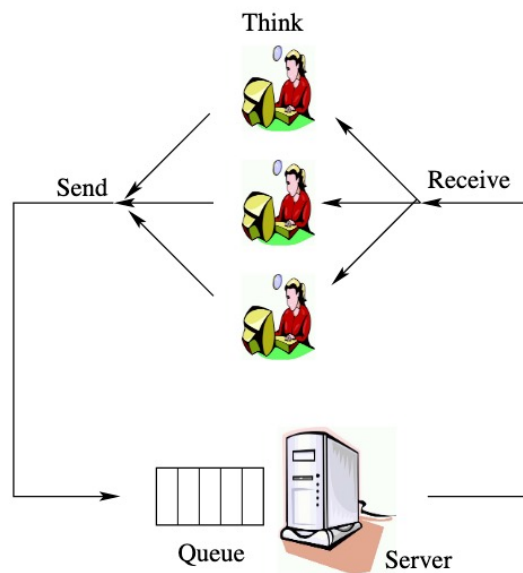
Research Roadmap 1:

- How do systems behave above saturation?
- How do those results compose?
- How do layers above know saturation is coming?

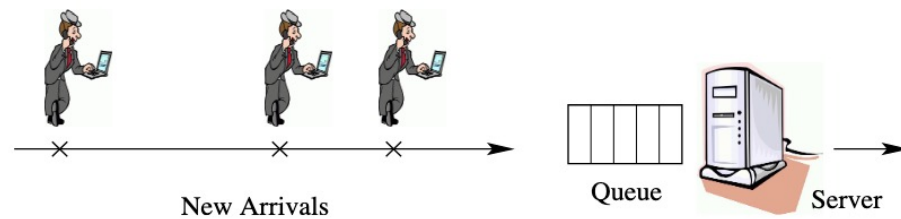
Why 2: Not Paying Attention to *Open* Workloads

Many benchmarks set a number of threads, then spin in a closed loop.

These benchmarks suck.



(a) Closed system

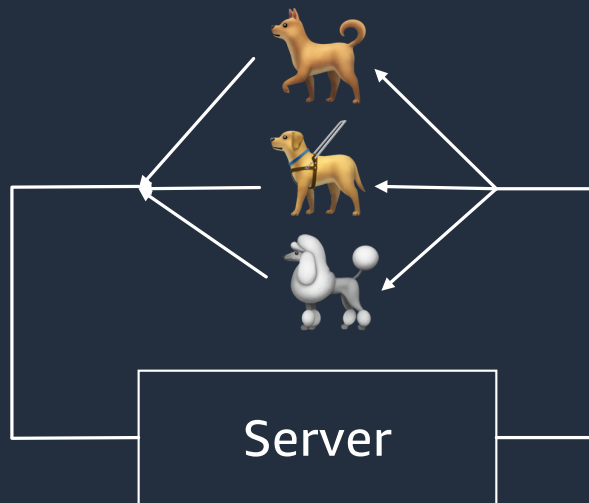


(b) Open system

From "Open Versus Closed: A Cautionary Tale", Schroeder et al, NSDI'06

They underestimate the effects of latency spikes.
(This is the famous *coordinated omission* problem).

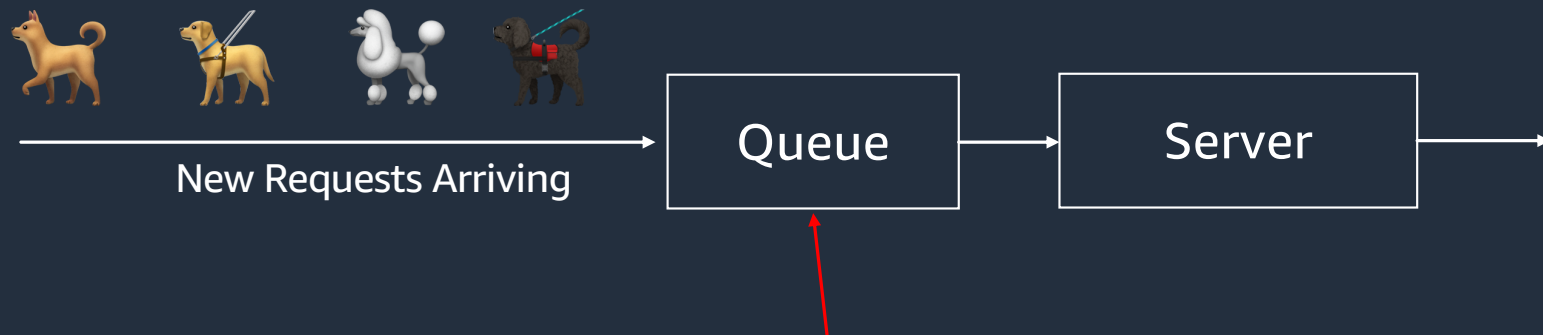
GOOD PERFORMANCE FOR BAD DAYS



Only one request per thread
observes a stall.

They're unrealistically kind.
When latency goes up, load goes down!

GOOD PERFORMANCE FOR BAD DAYS



Slowed down? Hope you can deal with more work later!

From "Open Versus Closed: A Cautionary Tale", Schroeder et al, NSDI'06



© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

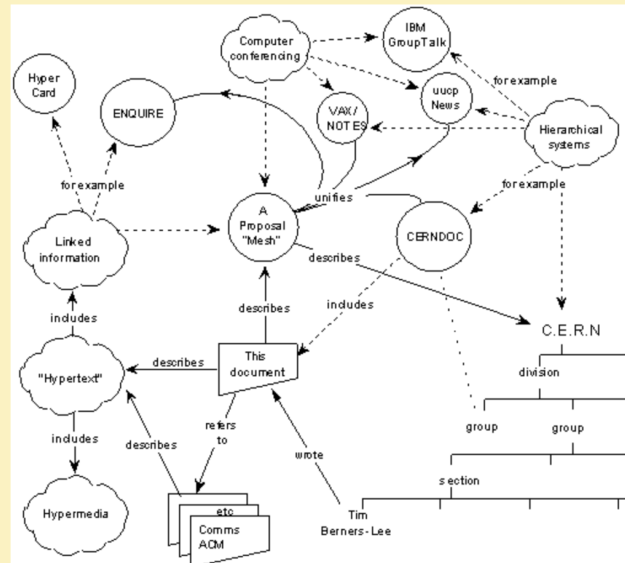
They don't reflect the reality most system builders live in.

Information Management: A Proposal

Tim Berners-Lee, CERN

March 1989, May 1990

This proposal concerns the management of general information about accelerators and experiments at CERN. It discusses the problems of loss of information about complex evolving systems and derives a solution based on a distributed hypertext system.



Overview

Many of the discussions of the future at CERN and the LHC era end with the question - "Yes, but how will we ever keep track of such a large project?" This proposal provides an answer to such questions. Firstly, it discusses the problem of information access at CERN. Then, it introduces the idea of linked information systems, and compares them with less flexible ways of finding information.

It then summarises my short experience with non-linear text systems known as "hypertext", describes what CERN needs from such a system, and what industry may provide. Finally, it suggests steps we should take to involve ourselves with hypertext now, so that individually and collectively we may understand what we are creating.

Research Roadmap 2:

- Benchmarks that represent 21st century workloads
- How does performance compose?
- How do clients and customers react to errors and latency?

Why 3: Not Paying Attention to *Metastability*

Overload conditions are *sticky*.

THE JOURNAL
of the
Acoustical Society of America

VOL. 31, NO. 1, JANUARY, 1959

4. Auditory Perception of Submerged Objects by Porpoises W. N. Kellogg 1

Reproduction and Identification of Elements of Auditory Displays Irwin Pollack and Lawrence B. Johnson 7

On Some Poststimulatory Effects at the Threshold of Audibility J. Zwislöcki, E. Picodda, and H. Rubin 9

9. Intelligibility of Peak-Clipped Speech at High Noise Levels . Irwin Pollack and J. M. Pickett 14

10. Scattering of High-Frequency Sound Waves in Polycrystalline Materials A. B. Bhatia 16

Optical Methods for the Measurement of the Sound Pressure in Liquids M. A. Breazeale and E. A. Hiedemann 24

Ultrasonic Absorption and Thermal Relaxation in CO₂ M. C. Henderson and J. Z. Klose 29

Propagation of Sound in Plate-Shaped Solid Delay Lines Paul M. Sutton 34

Diffraction of Light by Ultrasonic Waves Progressing with Finite but Moderate Amplitudes in Liquids K. L. Zankel and E. A. Hiedemann 44

11. Cavitation Microstreaming Samuel A. Elder 54

12. Statistical Evaluation of Near-Field Sound Pressures Generated by the Exhaust of a High-Performance Jet Engine Edwin C. Kamps 65

Theory of Noise and Vibration Insulation of a System with Many Resonances Eugen Skudrzyk 68

13. Tables of the Speed of Sound in Water Martin Greenspan and Carroll E. Tschiegg 75

2. American Standard Recommended Practice for Laboratory Measurement of Air-Borne Sound Transmission Loss of Building Floors and Walls K. U. Ingard, R. Hurdley, W. A. Jack, M. B. Summerfield, F. Tyzzer, and R. V. Waterhouse 76

On the Acoustics of Cocktail Parties William R. MacLean 79

Letters to the Editor:

9. Binaural Communication Systems: Preliminary Examination Irwin Pollack 81

Acoustical News 82

Current Publications on Acoustics:

Book Reviews 85

References to Contemporary Papers on Acoustics Walter Koidan and Earl D. Schubert 87

Review of Acoustical Patents Robert W. Young 99

Program of the Fifty-Sixth Meeting of the Acoustical Society of America 109

THE ACOUSTICAL SOCIETY OF AMERICA

ASA

1923

aws

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

24

13.	Tables of the Speed of Sound in Water	Martin Greenspan and Carroll E. Tschiegg	75
2.	American Standard Recommended Practice for Laboratory Measurement of Air-Borne Sound Transmission Loss of Building Floors and Walls	K. U. Ingard, R. Huntley, W. A. Jack, M. B. Summerfield, F. Tyzzer, and R. V. Waterhouse	76
	On the Acoustics of Cocktail Parties	William R. MacLean	79
<hr/>			
Letters to the Editor :			
9.	Binaural Communication Systems: Preliminary Examination	Irwin Pollack	81

Network Working Group

Request For Comments: 896

John Nagle

6 January 1984

Ford Aerospace and Communications Corporation

Congestion Control in IP/TCP Internetworks

This memo discusses some aspects of congestion control in IP/TCP Internetworks. It is intended to stimulate thought and further discussion of this topic. While some specific suggestions are made for improved congestion control implementation, this memo does not specify any standards.

“As long as there is only one copy of each datagram in transit, congestion is under control. Once retransmission of datagrams ... begins, there is potential for serious trouble.

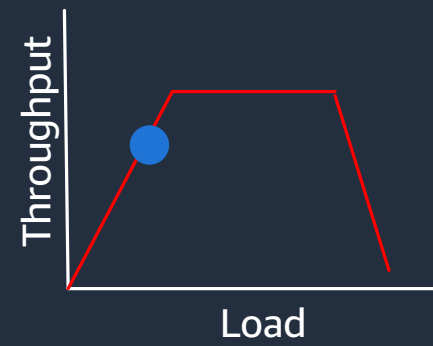
....

This condition is stable. Once the saturation point has been reached ... the network will continue to operate in a degraded condition.”

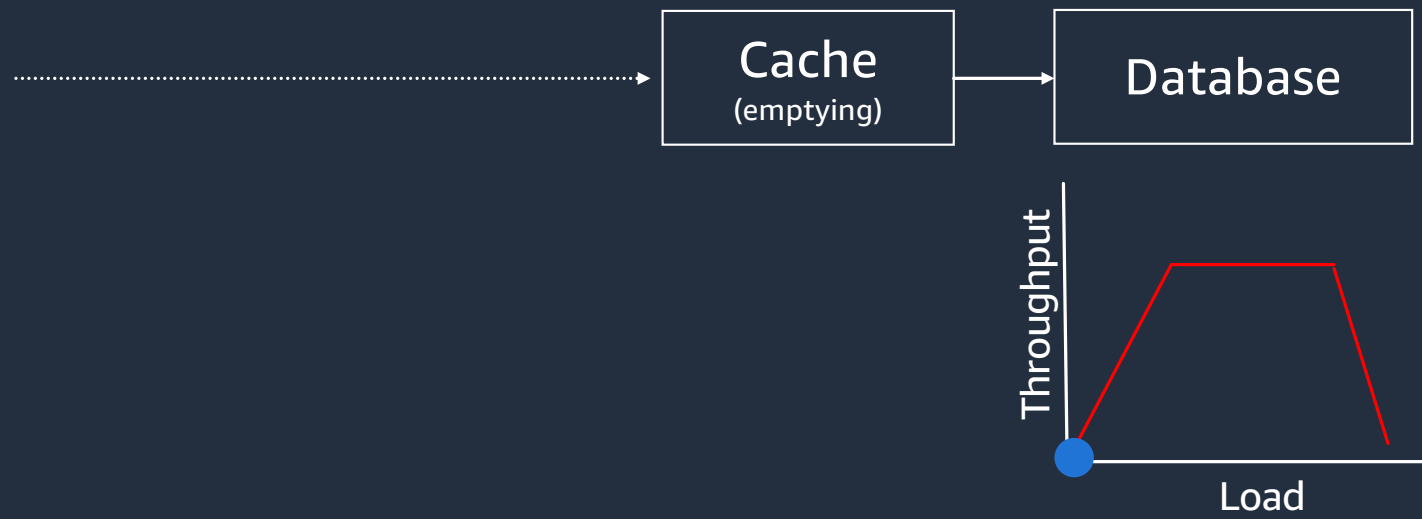
GOOD PERFORMANCE FOR BAD DAYS



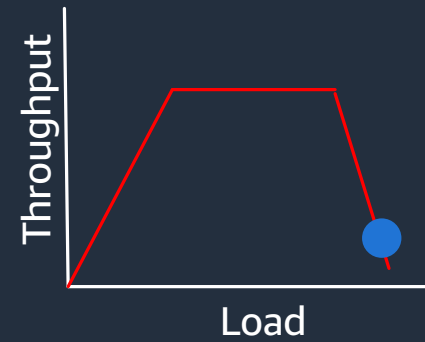
GOOD PERFORMANCE FOR BAD DAYS



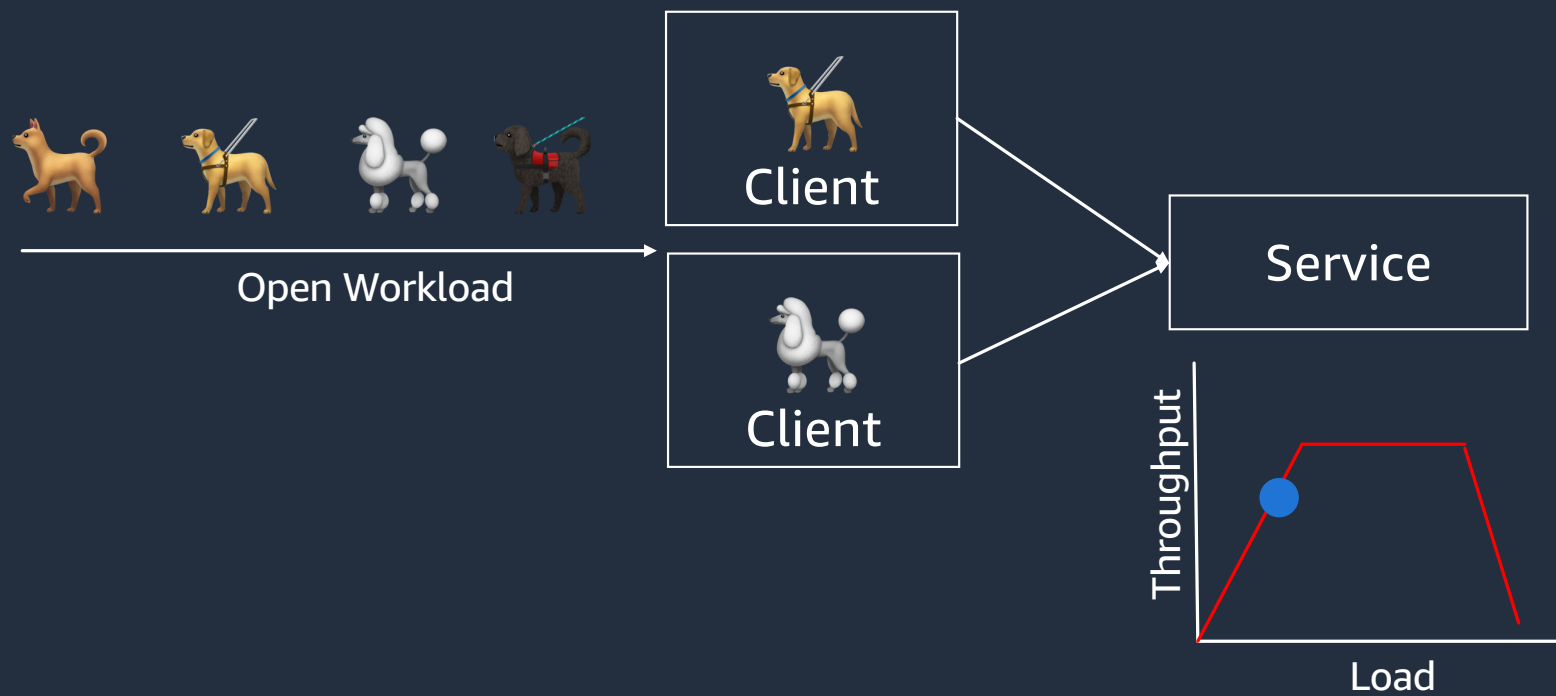
GOOD PERFORMANCE FOR BAD DAYS



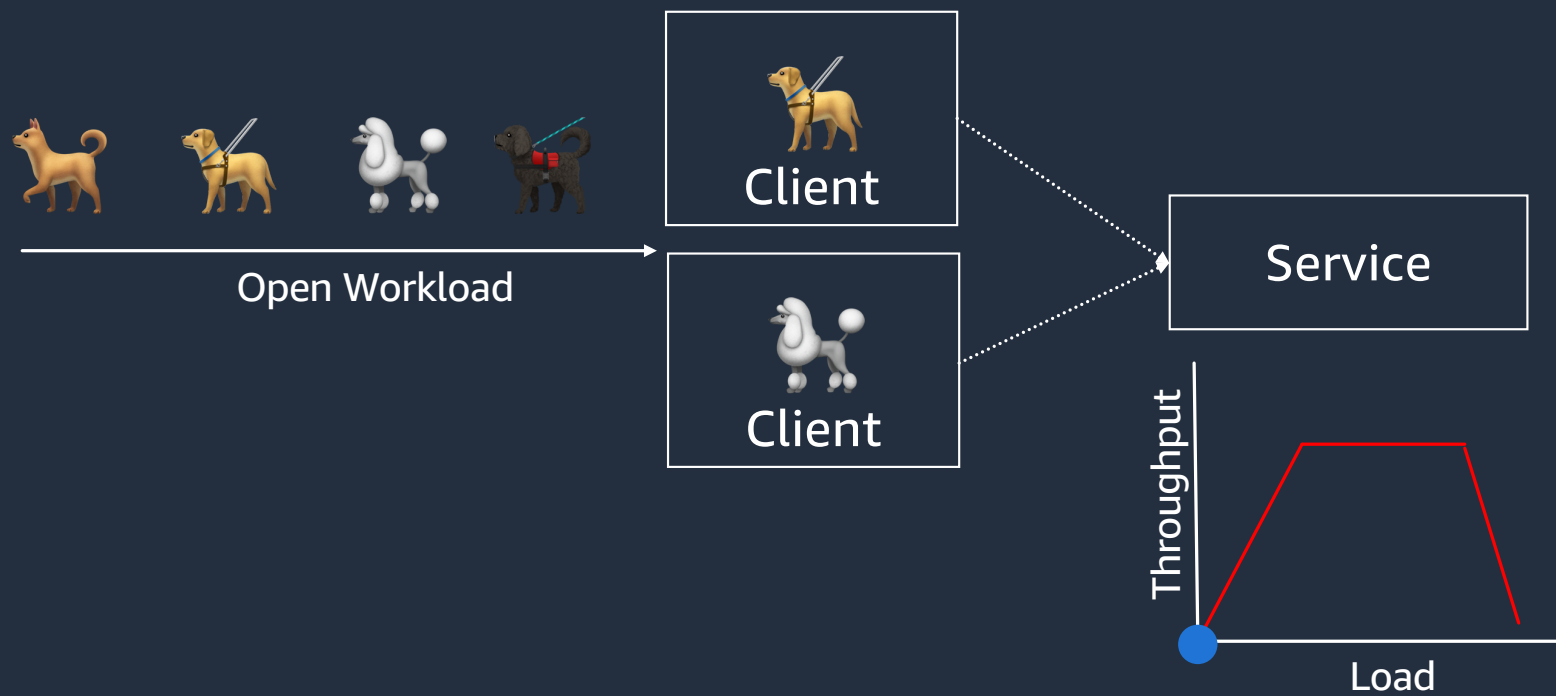
GOOD PERFORMANCE FOR BAD DAYS



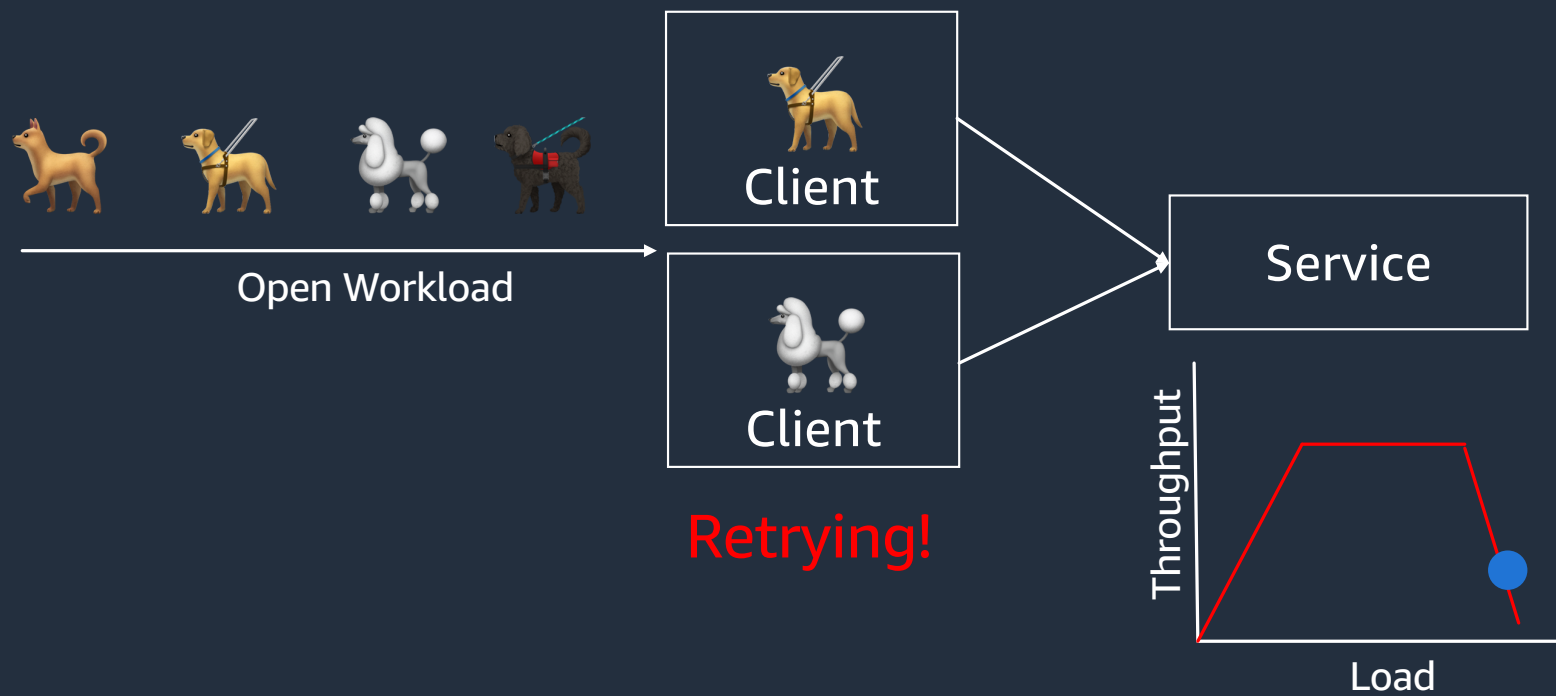
GOOD PERFORMANCE FOR BAD DAYS



GOOD PERFORMANCE FOR BAD DAYS



GOOD PERFORMANCE FOR BAD DAYS



Research Roadmap 3:

- Deeply understand metastable system behaviors.
- Formal and simulation tools to explore these behaviors *in silico*.
- Develop and advocate for better best practices.
- Algorithms for shedding load, detecting overload, etc.
- Incentives for client co-operation.

Why 4: Common Benchmarks Don't Pay Attention to *Workloads*

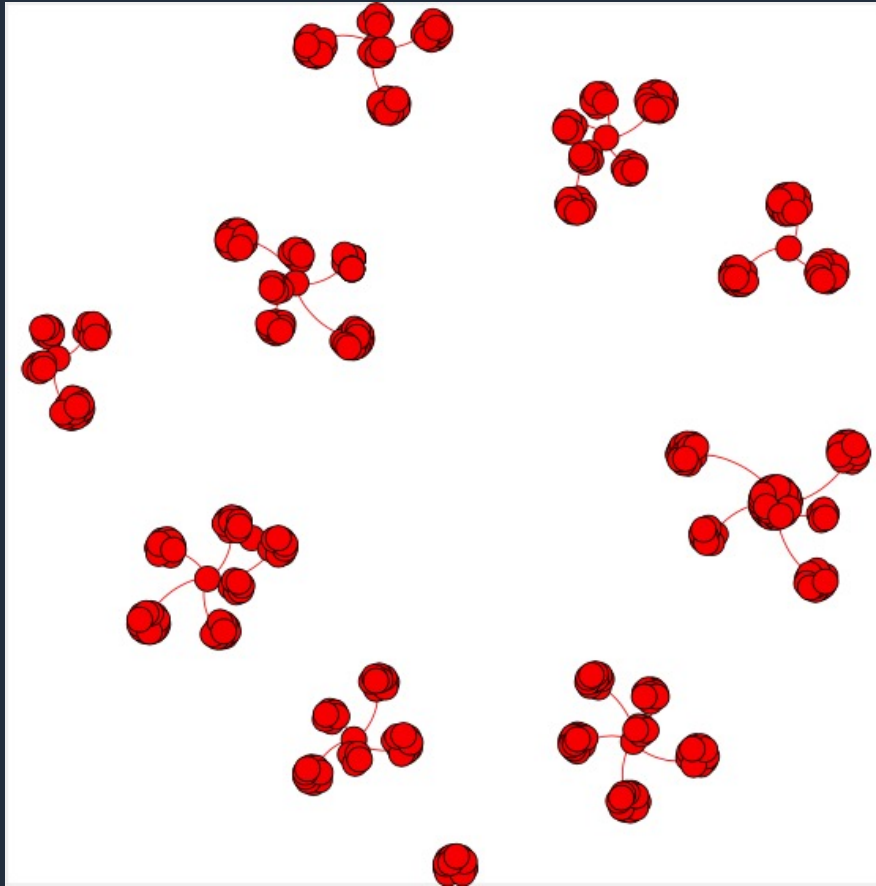
In many systems (especially RDBMSs), performance is very sensitive to workload.

This is well known, but not well used.

GOOD PERFORMANCE FOR BAD DAYS



GOOD PERFORMANCE FOR BAD DAYS



10 Warehouses
=
10 *Clusters*
=
Easy scaling

GOOD PERFORMANCE FOR BAD DAYS

Real world workloads are not as kind.



© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Table 3: Query Runtimes: *Number of queries and their total runtimes grouped into various runtime buckets.*

Time bucket	% of queries			% of sum(runtime)		
	Fleet	TPC-H	-DS	Fleet	TPC-H	-DS
(0s, 10ms]	13.7	0	0	0.01	0	0
(10ms, 100ms]	48.3	0	0	0.4	0	0
(100ms, 1s]	24.9	0	22	2.3	0	2
(1s, 10s]	9.9	27	59	7.3	3	19
(10s, 1min]	2.2	55	13	13.3	30	25
(1min, 10min]	0.86	8	5	35.7	66	55
(10min, 1h]	8e-2	0	0	25.2	0	0
(1h, 10h]	8e-3	0	0	14.3	0	0
>=10h	9e-5	0	0	1.6	0	0

Van Renen, et al. “Why TPC Is Not Enough: An Analysis of the Amazon Redshift Fleet”, VLDB’24

Research Roadmap 4:

- Characterization at the graph level.
- Focus on need for co-ordination, at fundamental level.
- Workload generation, and space exploration.



Thank you!

Marc Brooker

mbrooker@amazon.com

<https://brooker.co.za/blog/>

GOOD PERFORMANCE FOR BAD DAYS

